Elimination of a Misfolded Folding Intermediate by a Single Point Mutation

Jesper E. Mogensen,[‡] Henrik Ipsen,[§] Jens Holm,[§] and Daniel E. Otzen*,[‡]

Department of Life Sciences, Aalborg University, Sohngaardsholmsvej 49, DK-9000 Aalborg, and ALK-Abelló A/S, Bøge Alle 12-14, DK-2970 Hørsholm, Denmark

Received October 16, 2003; Revised Manuscript Received January 8, 2004

ABSTRACT: We present an analysis of the folding behavior of the 159-residue major birch pollen allergen Bet v 1. The protein contains a water-filled channel running through it. Consequently, the protein has a hydrophobic shell, rather than a hydrophobic core. During the folding of the protein from either the urea-, pH-, or SDS-denatured state, Bet v 1 transiently populates a partially folded intermediate state. This state appears to be misfolded, since it has to unfold at least partially to fold to the native state. The misfolded intermediate is not, however, a result of the water-filled channel in Bet v 1. The intermediate completely disappears in the mutant Tyr → Trp120, in which the channel is still present. Tyr120 appears to behave as a "negative gatekeeper" which attenuates efficient folding. The close structural homologue, the apple allergen Mal d 1, also folds without any detectable folding intermediates. However, the position of the transition state on the reaction coordinate, which is a measure of its overall compactness relative to the denatured and native states, is reduced dramatically from ca. 0.9 in Bet v 1 to around 0.5 in Mal d 1. We suggest that this large shift in the transition state structure is partly due to different local helix propensities. Given that individual mutations can have such large effects on folding, one should not *a priori* expect structurally homologous proteins to fold by the same mechanism.

When interest in the kinetics of protein folding started to grow in the early 1960s and early 1970s, the prevailing view was that proteins had to fold by specific "folding pathways" (*1*, *2*). That is, all protein molecules followed the same structural road map and folded to the native state through a series of well-defined states. In this *Weltanschauung*, intermediate or partially folded states were seen as key players whose structural properties, as well as the kinetics of their formation and decay, were central to the elucidation of the so-called folding code. Folding could be seen as a progression through a set of states with increasing degrees of native structure until the native state was finally attained. Once intermediates were observed experimentally, the question arose whether they were on- or off-pathway, that is, obligatory milestones leading in the right direction or kinetic traps which the protein had to escape by unfolding to get back on the "track" to the native state. There was kinetic evidence for both kinds of intermediates (*3*, *4*). The perceived role of intermediate states was altered by two developments in the 1990s. The first was the observation that many small proteins fold without detectable kinetic intermediates (*5*, *6*), indicating that intermediates were not essential in the search through conformational space. In cases when this class of proteins could be induced to form intermediates, for example, by performing the folding experiment in the presence of inorganic salts which favor compact states, these states were observed to slow folding in a manner consistent with an "off-pathway" intermediate (*7*). The second was the rise of a new view of folding, which replaced folding pathways with energy landscapes and folding funnels (*8*, *9*). In this view, there are many routes to the native state, and one cannot talk directly of on- versus off-pathway, although there may be preferred pathways (*10*). A smooth funnel with no major obstacles will lead to fast two-state kinetics, whereas in a more rugged landscape, the protein will experience a greater likelihood of being trapped temporarily in local minima, leading to multistate kinetics. In such a context, on- and off-pathway concepts lose their meaning, although one can still use terms such as "folding detours" and "misfolding" to indicate that the conformation of the protein in such minima contains elements of non-native structure which have to be unraveled for the protein to attain the native state. Note that although a folding process involves the entire energy landscape, experimentalists are limited by what they observe, i.e., the simple relaxation phases typically observed during folding and unfolding, which reflect the minima and maxima populated by protein states or ensembles in free energy profiles. This means that traditional pathway schemes are still the preferred medium for analyzing the experimental variables (*11*). Nevertheless, intermediate states retain their relevance for protein folding, since characterization of the properties of such partially folded states gives greater insight into the details of the energy landscape, and provides data for theoreticians to reproduce in folding simulations (*12*).

In this report, we describe the observation of a partially folded state during the folding of the major birch pollen allergen Bet v 1,[1] a 159-residue protein, consisting of a seven-stranded antiparallel β-sheet wrapped around a 25-residue C-terminal α-helix (Figure 1A) (*13*). Two small helices separate the β-sheet from the C-terminal end of the long helix. Bet v 1 contains a large, forked cavity, 30 Å long with a volume of approximately 1500 Å^3, which runs through the entire structure and has three openings to the protein surface. The surface of the cavity is mainly hydrophobic but

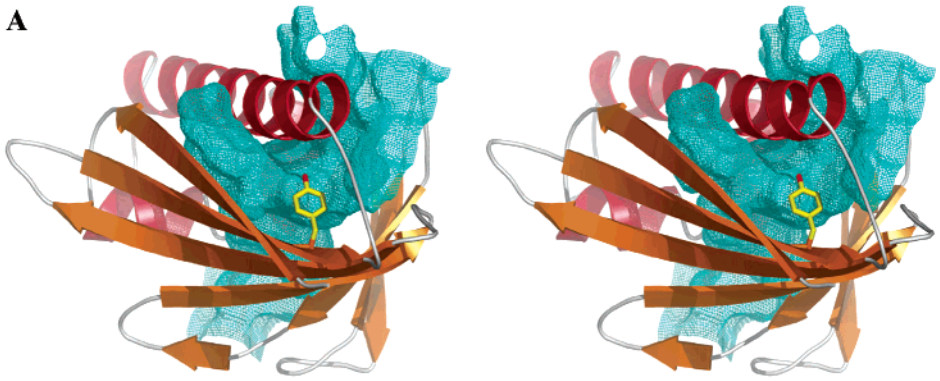
* To whom correspondence should be addressed. E-mail: dao@bio.auc.dk. Phone: (0045) 96358525. Fax: (0045) 98141808.

[‡] Aalborg University.

[§] ALK-Abelló A/S.

10.1021/bi0358622 CCC: $27.50
Published on Web 03/02/2004

A



B

```
                1                                                  50
Bet v 1 GVFNYETETT SVIPAARLFK AFILEGDTLI PKVAPQAISS ENIEGNGGP
Mal d 1 GVYTFENEFT SEIPPSRLFK AFVLDADNLI PKIAPQAIKQ EILEGNGGP
Consensus GV%n%EnEfT SeIPaaRLFK AF!L#aDnLI PK!APQAIkq EniEGNGGP
               51                                                 100
Bet v 1 GTIKKITFPE GSPFKYVKER VDEVDHANFK YSYSMIEGGA GDTLEKICN
Mal d 1 GTIKKITFGE GSQYGYVKHR IDSIDEASYS YSYTLIEGDA TDTIEKISY
Consensus GTIKKITFgE GSq%gYVKeR !De!DeAn%k YSYs$IEGdA gDTiEKIcn
               101                                                150
Bet v 1 EIKIVATPDG GSILKISNKY HTKGDHEMKA EHMKAIKEKG ALLRAVESY
Mal d 1 ETKLVACGSG STIKSISH.Y HTKGNIEIKE EHVKVGKEKA GLFKLIESY
Consensus EiKiVAcgdG gsIlkISn.Y HTKG#hEiKa EHmKagKEKa aLlra!ESY
               151       ;                ! is anyone of IV
Bet v 1 LLAHSDAYN ;                       $ is anyone of LM
Mal d 1 LKDHPDAYN ;                       % is anyone of FY
Consensus LlaHpDAYN ;                     # is anyone of NDQEBZ
```

FIGURE 1: (A) Stereoview of Bet v 1, indicating the position of the mutated Tyr120 and the cavity traversing the length of the molecule. The structure of the cavity was determined using the cavity calculation option in Voidoo (*60*). This probe-occupied cavity is defined by the surface of a rolling probe with a radius of 1.4 Å. The position of water molecule 263 was used as a starting point for the cavity calculation. This portion of the figure was made using PyMol (*61*). (B) Alignment of Bet v 1 and Mal d 1 sequences using Multalin (*62*).

contains polar residues D27, D69, Y81, Y83, N100, Q132, and S136, all of which hydrogen bond to four water molecules found in the cavity (*13*). Because of this cavity, Bet v 1 does not contain a dense, hydrophobic core and has a high surface area:size ratio, although it does not contain more polar residues than an average globular protein [47.7% for Bet v 1 vs an average of 46.2% (*14*)]. The cavity probably serves as a large ligand-binding pocket, and we have previously used an ANS displacement assay to identify possible physiological ligands such as fatty acids and plant hormones with K_d values in the low micromolar range (*15*). This has been corroborated by recent crystallographic and mass spectrometric studies of Bet v 1 in complex with deoxycholate and brassinosteroids (*16*). Deoxycholate binds to the cavity by a combination of H-bonds (to Bet v 1 backbone atoms and side chains) and extensive van der Waals contacts (between the polycyclic ring of deoxycholate and aromatic and aliphatic side chains lining the cavity) (*16*).

The study presented here was prompted by curiosity with regard to the folding of a protein lacking a large hydrophobic core. The hydrophobic effect, which is the entropically favorable release of water caused by the burial of hydrophobic surface area to form a hydrophobic core, is generally regarded as one of the central driving forces in protein folding (*17*), and its absence in Bet v 1 might be expected to lead to slow folding or multistate kinetics. Our data suggest that a partially folded state is formed transiently during Bet v 1's folding; however, this state appears to be a kinetic trap which is not a consequence of the lack of a hydrophobic core, since it is eliminated by a single mutation and does not form for a close structural homologue of Bet v 1. Thus, the folding landscape of Bet v 1 is rugged but is not significantly affected by the presence of the cavity.

[1] Abbreviations: ANS, 8-anilino-1-naphthalenesulfonic acid; Bet v 1, major birch allergen; C, off-pathway intermediate state; CD, circular dichroism; D, denatured state; I, on-pathway intermediate state; k_f, folding rate constant; k_{obs}, observed rate constant; k_u, unfolding rate constant; Mal d 1, major apple allergen; MES, 2-(*N*-morpholino)-ethanesulfonic acid; MOPS, 3-(*N*-morpholino)propanesulfonic acid; N, native state; SDS, sodium dodecyl sulfate; Tris, tris(hydroxymethyl)-aminomethane; Y120W, Tyr120 → Trp mutant of Bet v 1.

EXPERIMENTAL PROCEDURES

Materials. Factor Xa and the pMAL-c vector were from New England Biolabs (Beverly, MA). α-Cyclodextrin was from ICN Biomedicals (Aurora, OH). All other chemicals were from Sigma-Aldrich (St. Louis, MO). Bet v 1 isoform 2801 was used in this study. The Bet v 1 Y120W mutant was constructed using QuickChange site-directed mutagenesis (Stratagene, La Jolla, CA) using the manufacturer's instructions. The mutation and integrity of the open reading frame were confirmed by DNA sequencing. Wild-type Bet v 1, Bet v 1 Y120W, and Mal d 1 were expressed in fusions with the maltose binding protein using the pMAL-c vector and purified as described previously (*18*). Two other Bet v 1 mutants produced to report about structural changes elsewhere in the protein (F22W and F58W) were only obtained in very low yields and were not studied further.

Equilibrium Denaturation Experiments. Equilibrium fluorescence studies were carried out by excitation at 278 nm and emission at 308 nm (wild-type Bet v 1) or excitation at 280 nm and emission at 350 nm (Bet v 1 Y120W) on an RTC2000 spectrometer from Photon Technology International (Lawrenceville, NJ). Five scans were averaged to yield the final spectrum. Equilibrium circular dichroism studies were performed on a Jasco J-715 spectropolarimeter (Jasco Spectroscopic Co., Hachioji, Japan) with a Jasco PTC-348W1 temperature control unit. Six scans were averaged to yield the final spectrum.

Urea denaturation experiments were carried out at protein concentrations of ca. 2.5 μM (fluorescence studies on wild-type Bet v 1 and Y120W) or 10 μM (circular dichroism studies on Mal d 1 and wild-type Bet v 1) in 10 mM MOPS (pH 7) at 25 °C. Urea stock solutions (10 M) were prepared fresh on a daily basis. Each protein sample was allowed to equilibrate for 2 h before being measured. Ellipticity values at 222 nm were used, since they showed the greatest change upon denaturation. Absorption by urea prevented us from going below 215 nm. The ellipticity and fluorescence values are analyzed using the linear extrapolation method (*19*, *20*), which is based on the relationship[2]

$$\log K_{\mathrm{D-N}}{}^{\mathrm{urea}} = \log K_{\mathrm{D-N}}{}^{\mathrm{water}} + m_{\mathrm{D-N}}[\mathrm{urea}] \qquad (1)$$

where $K_{\mathrm{D-N}}$ ([D]/[N]) is the equilibrium constant of unfolding and $m_{\mathrm{D-N}}$ is a parameter which reflects the degree of surface area buried in the native state, N, relative to the denatured state, D (*19*). The larger the *m* value, the greater the difference between N and D in exposed surface area. The advantage of this standard linear extrapolation approach is that it allows us to compare equilibrium and kinetic data, and thus test the consistency of kinetic interpretations (see below). Assuming a linear dependence of the pre- and post-transition baselines on urea concentration (*20*, *21*), we obtain the following equation for urea denaturation curves:

$$Y_{\mathrm{obs}} = \frac{\alpha_{\mathrm{N}} + \beta_{\mathrm{N}}[\mathrm{urea}] + (\alpha_{\mathrm{D}} + \beta_{\mathrm{D}}[\mathrm{urea}]) \times 10^{m_{\mathrm{D-N}}([\mathrm{urea}]-[\mathrm{urea}_{50\%}])}}{1 + 10^{m_{\mathrm{D-N}}([\mathrm{urea}]-[\mathrm{urea}_{50\%}])}} \qquad (2)$$

where Y_{obs} is the observed signal, α_{N} and α_{D} denote the signal at 0 M urea for the native and denatured states, respectively, β_{N} and β_{D} are the slopes of the baselines of the native and denatured states, respectively, and $[\mathrm{urea}_{50\%}]$ is the urea concentration where 50% of the protein is denatured. Nonlinear least-squares regression analysis was carried out with Kaleidagraph, version 3.5 (Synergy Software, Reading, PA). From the fit to eq 2, the stability of Bet v 1 in water is calculated as follows (*20*, *21*):

$$\Delta G_{\mathrm{D-N}} = -RT \ln(10) m_{\mathrm{D-N}}[\mathrm{urea}_{50\%}] \qquad (3)$$

pH-induced denaturation experiments were performed using CD at a protein concentration of 10 μM. The ellipticity was followed at 222 nm in the following buffers (50 mM): 0.1 M HCl at pH 1, glycine at pH 2−3.5, sodium acetate at pH 3.7−4.7, MES at pH 6, MOPS at pH 7, Tris at pH 8−9, glycine at pH 10−11, and 0.1 M NaOH at pH 13. Thermal scans at different pH values were carried out using these buffers.

Thermal CD denaturation experiments were carried out at protein concentrations of 10 μM in 50 mM buffer with a scan rate of 60 °C/h at 220 nm. Data were fitted to the following equation (*22*):

$$\theta = \frac{\alpha_{\mathrm{N}} + \beta_{\mathrm{N}}T + (\alpha_{\mathrm{D}} + \beta_{\mathrm{D}}T)\mathrm{e}^{-[\Delta H_{T\mathrm{m}}(1-T/T_{\mathrm{m}})]/RT}}{1 + \mathrm{e}^{-[\Delta H_{T\mathrm{m}}(1-T/T_{\mathrm{m}})]/RT}} \qquad (4)$$

where θ is the raw CD signal, α_{N} and α_{D} are the θ values for the folded and denatured states, respectively, at 298 K, β_{N} and β_{D} are the slopes of the baselines of the native and denatured states, respectively, and $\Delta H_{T_{\mathrm{m}}}$ is the enthalpy of unfolding at the midpoint of denaturation (T_{m}). When $\Delta H_{T_{\mathrm{m}}}$ and T_{m} are measured at different pH values, the specific heat capacity ΔC_p can be obtained as the slope of the plot of $\Delta H_{T_{\mathrm{m}}}$ versus T_{m} (assuming for simplicity that ΔC_p is constant in the experimental temperature range). Using this value of ΔC_p, the stability of Bet v 1 at any temperature can be calculated as

$$\Delta G_T = \Delta H_T - T\Delta S_T = \Delta H_{T_{\mathrm{m}}}\left(1 - \frac{T}{T_{\mathrm{m}}}\right) - \Delta C_p\left[T_{\mathrm{m}} - T + T\ln\left(\frac{T}{T_{\mathrm{m}}}\right)\right] \qquad (5)$$

Stopped-Flow Folding and Unfolding Experiments. Stopped-flow fluorescence measurements were performed on an Applied Photophysics (Leatherhead, U.K.) SX18MV stopped-flow apparatus with a 305 nm (wild-type Bet v 1) or 320 nm (Bet v 1 Y120W) glass filter. Folding and unfolding experiments were initiated by 10-fold dilution of the protein from 4.4 and 0 M urea, respectively, to the appropriate final urea concentrations in 10 mM MOPS (pH 7). This allowed us to measure over the range of 0.4−7 M urea. To assess refolding at lower urea concentrations, pH-jump experiments were carried out. The protein was denatured at pH 2.1 (where unfolding is most pronounced; see Figure 3A) in HCl and refolded into final urea concentrations of 0−0.6 M in 50 mM MOPS (pH 7). Refolding from the SDS-denatured state was carried out by denaturing the protein in 14 mM SDS

[2] To simplify comparison between equilibrium and kinetic data, we use log $K_{\mathrm{D-N}}$ rather than $\Delta G_{\mathrm{D-N}}$. The two parameters are interconverted with the relationship $\Delta G_{\mathrm{D-N}} = -1.36 \log K_{\mathrm{D-N}}$.

and diluting it 6-fold into 16 mM α-cyclodextrin in 10 mM MOPS (pH 7) (*23*).

For stopped-flow CD measurements, the SFM-400 accessory unit from Biologic (Claix, France) was mounted on the Jasco J-715 CD spectropolarimeter. On average, 12 runs were recorded and the subsequent time profile was fitted to a single-exponential function. Folding and unfolding experiments (Mal d 1 and wild-type Bet v 1) were carried out by 4-fold dilution of the protein into the appropriate buffer. Refolding experiments were here carried out from the pH-denatured state.

The analysis of the kinetic data is based on a linear relationship between the log of microscopic rate constants and the urea concentration (*5*, *19*), in a manner similar to eq 1:

$$\log k_x^{\text{urea}} = \log k_x^{\text{water}} + m_x[\text{urea}] \quad (6)$$

where x refers to folding (f) or unfolding (u) and the *m* values refer to the burial of surface area in the transition state relative to the ground state from which the reaction starts. In the simple two-state case, where there is no intermediate between D and N, the observed rate constant $k_{obs} = k_f + k_u$. As a consequence of this, the chevron plot (a plot of log k_{obs} vs urea concentration) will be V-shaped, with two straight lines meeting at the midpoint of denaturation, $[\text{urea}]_{50\%}$ (cf. eq 7 below). A further test of the two-state model is the fact that $K_{D-N} = k_f/k_u$ and $m_{D-N} = m_f - m_u$. The accumulation of on- or off-pathway intermediates during folding will give rise to deviations from linearity in the refolding region of the plot (eqs 9 and 10). The above interpretation of the *m* values means that the m_f/m_{D-N} ratio (also known as β_T) is an indication of the compactness of the transition state of folding relative to the native state.

The following equations were used to fit rate constants determined from stopped-flow experiments (see Results for further details).

(A) Two-state folding according to (*5*)

Scheme 1

$$D \underset{k_u}{\overset{k_f}{\rightleftharpoons}} N$$

$$\log k_{obs} = \log[10^{(\log k_f^{water} + m_f[\text{urea}])} + 10^{(\log k_u^{water} + m_u[\text{urea}])}] \quad (7)$$

Here the observed rate constant $k_{obs} = k_f + k_u$.

(B) Two-state folding with curvature in the unfolding limb (Scheme 1 with a moving transition state) (*24*):

$$\log k_{obs} = \log[10^{(\log k_f^{water} + m_f[\text{urea}])} + 10^{(\log k_u^{water} + m_u[\text{urea}] + m_u^*[\text{urea}]^2)}] \quad (8)$$

Again, $k_{obs} = k_f + k_u$. m_u* is a constant which describes the curvature in the plot of log k_u versus urea concentration.

[3] Urea is used rather than the more potent denaturant guanidinium chloride for two reasons. First, denaturation occurs at low urea concentrations, and the use of guanidinium chloride leads to an even narrower concentration range for folding (data not shown). Second, urea is not a salt, and therefore, we avoid complications from changes in ionic strength.

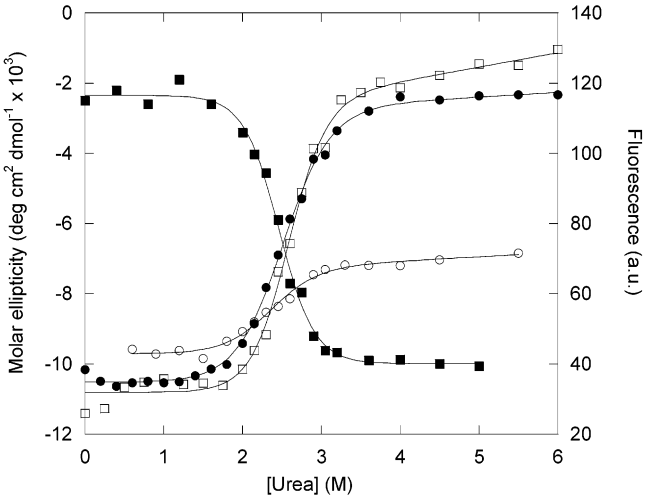


FIGURE 2: Equilibrium denaturation of wild-type Bet v 1 followed by CD (●) and fluorescence (○), of Bet v 1 Y120W followed by fluorescence (■), and of Mal d 1 followed by CD (□). The solid lines represent the best fits to eq 2.

(C) Folding with an on-pathway folding intermediate according to (*25*)

Scheme 2

$$D \underset{}{\overset{K_I}{\rightleftharpoons}} I \underset{k_u}{\overset{k_f}{\rightleftharpoons}} N$$

$$\log k_{obs} = \log\left[\frac{10^{(\log k_f^{water} + m_f[\text{urea}])}}{1 + 10^{-(\log K_I^{water} + m_I[\text{urea}])}} + 10^{(\log k_u^{water} + m_u[\text{urea}])}\right] \quad (9)$$

Here $k_{obs} = k_f f_I + k_u$, where $f_I = [I]/([I] + [D])$ and $K_I = [I]/[D]$.

(D) Folding with an off-pathway folding intermediate according to (*7*, *25*)

Scheme 3

$$C \underset{}{\overset{K_C}{\rightleftharpoons}} D \underset{k_u}{\overset{k_f}{\rightleftharpoons}} N$$

$$\log k_{obs} = \log\left[\frac{10^{(\log k_f^{water} + m_f[\text{urea}])}}{1 + 10^{(\log K_C^{water} + m_C[\text{urea}])}} + 10^{(\log k_u^{water} + m_u[\text{urea}])}\right] \quad (10)$$

Here $k_{obs} = k_f f_D + k_u$, where $f_D = [D]/([D] + [C])$ and $K_C = [C]/[D]$.

RESULTS

Equilibrium Denaturation: Bet v 1 Unfolds in a Cooperative Fashion and Exhibits Low Stability. We start out with a characterization of the equilibrium denaturation of Bet v 1 in urea so we can have a basis for comparing the kinetic data.[3] Bet v 1 unfolds with a single transition around 2.4 M urea according to both Tyr-based fluorescence spectroscopy and far-UV CD (Figure 2). The close correspondence between fluorescence and CD experiments indicates that folding intermediates do not accumulate to any significant extent under equilibrium conditions. The calculated stability of 5.84 ± 0.75 kcal/mol (Table 1) is relatively low, as might be expected for a protein with a reduced hydrophobic core.

Table 1: Kinetic Parameters for Folding of Wild-Type Bet v 1, Bet v 1 Y120W, and Mal d 1[a]

	wild-type Bet v 1						
	kinetics			Bet v 1 Y120W		wild-type Mal d 1	
parameter	on-pathway[b]	off-pathway[c]	equilibrium	kinetics[d]	equilibrium	kinetics[e]	equilibrium
m_f (M^{-1})	-0.013 ± 0.15	-1.79 ± 0.14	–	-1.59 ± 0.05	–	-1.02 ± 0.03	–
m_u (M^{-1})	0.23 ± 0.02	0.23 ± 0.02	–	0.27 ± 0.02	–	1.52 ± 0.14	–
m^* (M^{-2})	–[j]	–[j]	–	–	–	-0.10 ± 0.014	–
$m_u^{midpoint}$ (M^{-1})	–	–	–	–	–	1.06 ± 0.14[f]	–
m_l (M^{-1})	-1.78 ± 0.11	-1.78 ± 0.04	–	–	–	–	–
m_{D-N} (M^{-1})	2.01 ± 0.18	2.02 ± 0.14	1.77 ± 0.22	1.86 ± 0.04	1.93 ± 0.19	2.08 ± 0.14	2.04 ± 0.22
k_f (s^{-1})	12.3 ± 0.45	644 ± 62	–	125 ± 3.3	–	6.53 ± 0.26	–
k_u (s^{-1})	0.009 ± 0.0004	0.009 ± 0.0004	–	0.002 ± 0.0002	–	$(7.5 \pm 0.4) \times 10^{-5}$[g]	–
K_I (s^{-1})[c]	30 ± 4.9	51 ± 8	–	–	–	–	–
$[urea]_{50\%}$ (M)	2.51	2.51	2.43 ± 0.07	2.59 ± 0.09	2.46 ± 0.03	2.38	2.61 ± 0.03
ΔG_{D-N} (kcal/mol)[h]	-6.63 ± 0.40	-6.63 ± 0.40	-5.84 ± 0.75	-6.54 ± 0.16	-6.46 ± 0.64	-6.72 ± 0.31	-7.29 ± 0.79
β_T^{TS}[i]	0.87	0.89		0.86		0.49	

[a] All conditions at pH 7.0 and 25 °C. Errors are standard errors of the parameters from the best fits to the different equations. [b] Data based on the on-pathway folding scheme (D ↔ I ↔ N) (eq 9). [c] Data based on the off-pathway folding scheme (C ↔ D ↔ N) (eq 10). [d] Data based on a two-state folding scheme (D ↔ N) (eq 7). [e] Data based on two-state folding with curvature in the unfolding limb (D ↔ N) (eq 8). [f] $m_u^{midpoint} = m_u + 2[urea]_{50\%}m^*$. [g] Based on linear extrapolation from the midpoint of denaturation. [h] Calculated from eq 3. [i] $\beta_T^{TS} = -m_f/(m_u^{midpoint} - m_f)$. [j] Unfolding data for Bet v 1 showed slight curvature and could be fitted to a second-order polynomium wih an m^* value of -0.031 ± 0.02 M^{-2}. This increased m_{D-N} from 2.01 to 2.15 ± 0.20 at the midpoint of denaturation, which does not significantly affect our conclusions.

Typically, globular proteins have ΔG_{D-N} values in the range of 5−15 kcal/mol (*26*).

To measure the stability of Bet v 1 by an independent method, we have used data from thermal denaturation scans. To calculate the stability of Bet v 1 at 25 °C, we need to determine the specific heat capacity (ΔC_p). Although ΔC_p values can in principle be obtained from individual thermal scans, these values are second derivatives of the temperature plot, making them notoriously inaccurate. Instead, we exploit the fact that $\Delta C_p = d\Delta H/dT$. Making the common assumption that ΔC_p is pH- and temperature-independent in the temperature range of 20−80 °C, one can determine ΔC_p from the slope of a plot of ΔH_{T_m} versus the melting temperature T_m (*27*). We performed thermal denaturation scans of Bet v 1 in the pH range of 3.5−10.1, where it is predominantly native at 25 °C (the pH denaturation plot shown in Figure 3A indicates an acid transition around pH 3.6 and an alkaline transition around pH 10.5; the protein has the highest T_m at pH 7, where it melts at 67.7 °C, as shown in Figure 3B). A plot of ΔH_{T_m} versus the melting temperature T_m (Figure 4) fits reasonably well to a straight line with a slope (ΔC_p) of 1.20 ± 0.19 kcal mol^{-1} K^{-1}. Using this value of ΔC_p in eq 5 together with the ΔH_{T_m} value intrapolated to 67.7 °C from Figure 4 (64.4 kcal/mol), we can calculate the stability of Bet v 1 at pH 7 using eq 5 to be 4.49 ± 0.75 kcal/mol, which is compatible with the urea denaturation data.

m_{D-N} and ΔC_p are both measures of the difference in solvent-accessible surface area between D and N. m_{D-N} reflects the difference in the number of unspecific denaturant binding sites, while ΔC_p predominantly comes from the release of water as nonpolar groups are buried (*28*−*30*), with the nuance that heat capacity changes for aliphatic and aromatic groups differ slightly (*28*, *30*). Since heat capacity changes and *m* values both reflect the burial of surface area, they should roughly correspond; indeed, for the cold-shock protein CspB, 90% of the total heat capacity change and 96% of the *m* value change occur between the unfolded state and the transition state (*31*). In the case of Bet v 1, the m_{D-N} value of 1.77 ± 0.22 M^{-1} is slightly higher than the corresponding values of most globular proteins with a similar number of residues (data not shown), whereas the ΔC_p value of 1.20 kcal mol^{-1} K^{-1} is rather low; a protein of 159 residues would be expected to have a value around 2.4 kcal mol^{-1} K^{-1} (*29*, *32*). This suggests that factors other than simple burial of hydrophobic surface area contribute to the magnitude of the m_{D-N} and ΔC_p values. For example, the desolvation of polar groups makes a small but positive contribution to ΔC_p (*28*), while the effect of these groups on the *m* value is unclear. Furthermore, the configurational freedom gained upon protein unfolding accounts for up to 20% of the total heat capacity increase (*33*).

Wild-Type Bet v 1 Folds According to a Three-State Model with an Intermediate that Has Off-Pathway Characteristics. Although Bet v 1 has no Trp residues, the seven Tyr residues give rise to a small but measurable signal change upon folding and unfolding (cf. Figure 2). We were therefore able to perform fluorescence-based stopped-flow experiments to measure refolding and unfolding rates (Figure 5). Unfolding data could be fitted to a single-exponential decay with a slow linear drift, the latter of which may be ascribed to photobleaching. Refolding rates, on the other hand, gave rise to clear double-exponential decays with two well-separated phases. Like the fast phase, the rate of the slow refolding phase decreased linearly with urea concentration (data not shown). It is customary to ascribe slow refolding phases to one of three phenomena, either transient aggregation during folding (*34*, *35*), the formation or decay of folding intermediates (*11*), or heterogeneity of the denatured state, commonly *cis*−*trans* isomerization about prolyl−peptidyl bonds (*36*−*38*). Aggregation can be ruled out, as the refolding rate constant did not vary significantly with concentration over a 10-fold range (data not shown). As described below, rate constants from the fast phase alone provide data for a model which is consistent with equilibrium denaturation data. This makes it possible to discount intermediate formation as the reason for the slow phase. This leaves *cis*−*trans* isomerization as the only remaining explanation. However, including the prolyl−peptidyl isomerase cyclophilin in the refolding buffer did not affect the slow rate constant when refolding from the acid-denatured state (data not shown). Thus, either

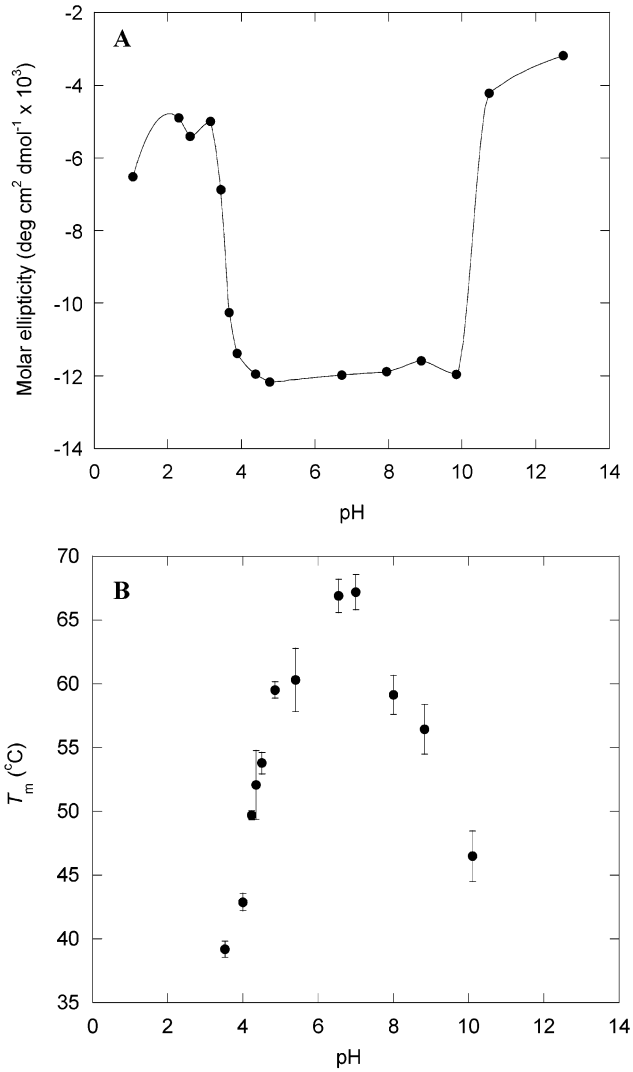


FIGURE 3: pH stability of wild-type Bet v 1 as measured by (A) ellipticity at 222 nm and (B) T_m determined from thermal scans followed by CD. The error bars represent the standard error of the mean from three experiments.

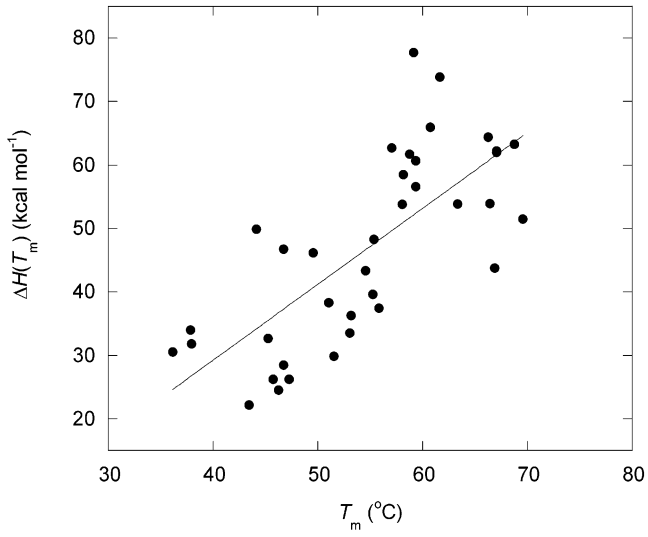


FIGURE 4: Enthalpy of unfolding (ΔH_{T_m}) for wild-type Bet v 1 vs T_m. Experiments were performed between pH 3.5 and 10.1. The slope (1.20 ± 0.19 kcal mol^{-1} K^{-1}) represents the specific heat capacity (ΔC_p).

traditional *cis*−*trans* proline isomerization in the denatured state is not rate-limiting, or cyclophilin does not have access to *cis* prolyl−peptidyl bonds in denatured Bet v 1. The latter could happen if the *cis* bonds impart more persistent structure to the denatured state. On the other hand, the slow refolding phase extrapolates to 0.3 s^{-1} in 0 M urea, which is rather fast for conventional isomerization steps [typical rate constants are on the order of 10^{-2}−10^{-4} s^{-1} (*39*)]. Nonprolyl *cis*−*trans* isomerization cannot be ruled out (*40*). In the following, we will concentrate on the fast phase only to analyze our data according to the "pathway" approach outlined in the introductory section and Experimental Procedures. Here pathway refers to the average trajectory of the whole interconverting protein ensemble through the energy landscape.

Folding experiments in which Bet v 1 is refolded by diluting out from the urea-denatured state (here 4.4 M) only allow us to assess refolding down to 0.4 M urea. It is important to go down to 0 M urea to investigate the possible existence of kinetic intermediates, because they tend to accumulate only at very low denaturant concentrations. To extend the range down to 0 M urea, we performed additional denaturation experiments in two different ways. (1) Bet v 1 was acid-denatured and refolded by pH-jumping it to pH 7 in the presence of 0−0.6 M urea. (2) Bet v 1 was denatured in SDS, and refolding was initiated by mixing it with excess α-cyclodextrin, which strips the SDS molecules off the protein, leaving the denatured protein to refold in the absence of SDS and urea. The latter method has been shown to lead to the expected folding rates for a number of different proteins, including ones with refolding rate constants in excess of 100 s^{-1} (*23*). There is good correspondence between rate constants for refolding from the acid- and urea-denatured state in the overlapping concentration region (Figure 5). In addition, the value for refolding from the SDS-denatured state in water (12.1 s^{-1}) agrees well with the value for the acid-denatured state measured by fluorescence (13.1 s^{-1}) and CD (10.1 s^{-1}).

Importantly, the combined refolding data clearly indicate that log k_f does not vary linearly with urea concentration in the whole refolding region of the chevron plot (Figure 5). Simple two-state folding predicts that both log k_f and log k_u vary linearly with urea concentration (eq 7 in Experimental Procedures). The so-called rollover below ca. 1 M urea, in which the slope of the refolding limb decreases significantly, suggests that an intermediate accumulates transiently during refolding. The rollover arises because of a switch in the ground state from which folding occurs, namely, from the intermediate state at low urea concentrations to the denatured state at higher concentrations (*25*). As the intermediate becomes more and more stable at lower denaturant concentrations, the activation barrier to refolding increases, slowing folding. However, on the basis of our data in Figure 5 alone, we cannot distinguish between productive on-pathway intermediate I (D ↔ I ↔ N, Scheme 2) and off-pathway intermediate C (C ↔ D ↔ N, Scheme 3) (*25*). In the latter case, the protein has to unfold from the C state back to the D state to fold. Both pathways can be fitted satisfactorily to the kinetic data and yield identical predictions for the stability of the native state (Table 1). Complete resolution of the pathway requires us to measure the rate constants for formation and decay of the intermediate (*41*), which is

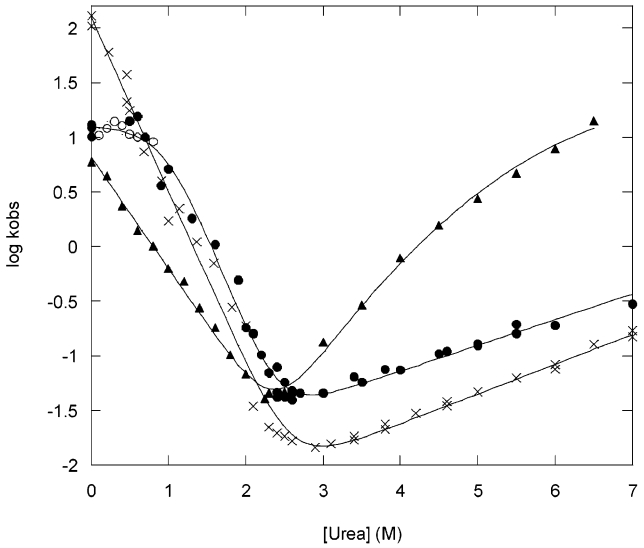


FIGURE 5: Log of the observed rate of folding and unfolding of wild-type Bet v 1 (○ and ●, respectively), of Bet v 1 Y120W (×), and of the structural homologue Mal d 1 (▲) vs urea concentration. In the refolding limb of wild-type Bet v 1, the empty circles represent refolding from the pH-denatured state and the filled circles that from the urea-denatured state except for the filled circle values at 0 M urea, which represent refolding rates from (a) the SDS-denatured state and (b) as measured by CD instead of fluorescence. The data for wild-type Bet v 1 can be fitted to a three-state scheme with an on- or off-pathway intermediate (eqs 9 and 10), while those for Y120W are fitted to a simple two-state folding scheme (eq 7) and those of Mal d 1 to a two-state folding scheme with unfolding curvature (eq 8).

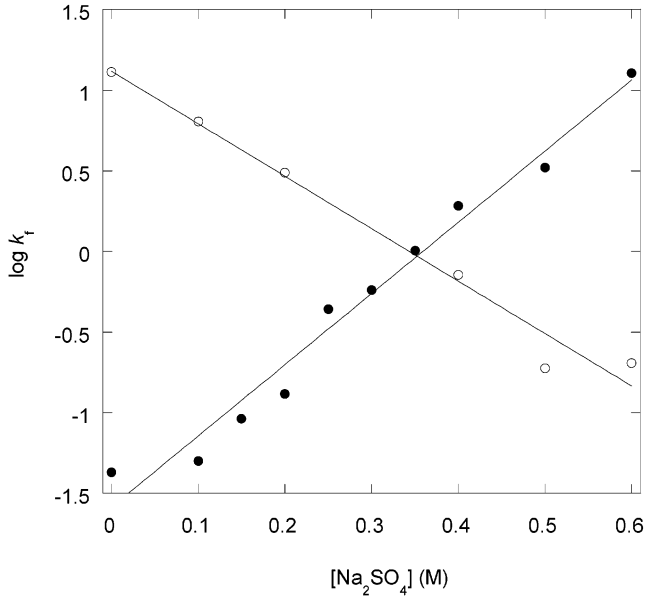


FIGURE 6: Log of the observed rate of folding of wild-type Bet v 1 vs Na_2SO_4 concentration in 0 M urea (○) and 2.1 M urea (●). In 0 M urea, the intermediate is the ground state from which refolding occurs, whereas in 2.1 M urea, folding occurs from the denatured state. The dramatic decrease in refolding rates with increasing salt concentrations at 0 M urea is strong evidence that the intermediate is off-pathway.

beyond our technical scope. Nevertheless, there is a simple assay which can shed more light on this. Inorganic salts such as sodium sulfate favor compact protein states because they are preferentially excluded from the protein surface (*42*). The more compact the state, the less salt must be excluded from the surface and the smaller entropic penalty the system must pay. Since a partially folded state is more compact than the denatured state, sodium sulfate will stabilize this state relative to the denatured state and at sufficiently high salt concentrations will force it to accumulate transiently during the refolding process. The native state will ultimately form, because it is the most compact state and therefore most favored by salt. However, even transient accumulation of a folding intermediate will have a profoundly different effect on the observed folding rate constant, depending on whether it is on-pathway or off-pathway. In the off-pathway scenario, folding is limited by the fraction of the protein which is in denatured state D [$f_D = [D]/([D] + [C])$], and the more C is stabilized by salts relative to D, the slower the observed rate constant. In contrast, the on-pathway scenario leads to folding being limited by the fraction f_I [$[I]/([I] + [D])$], and the more I is stabilized, the faster the rate constant should be (the transition state for folding, which is generally more compact than I, will be stabilized even more than I, and therefore, the activation barrier decreases with salt concentration). Under conditions where folding occurs directly from the denatured state and the intermediate does not accumulate, rate constants should also be accelerated by salt due to the stabilization of the transition state.

We have previously shown that the intermediate, which accumulates when ribosomal protein S6 folds in the presence of Na_2SO_4, is off-pathway, since it slows folding (*7*). In the case of Bet v 1, we have measured the effect of salt under two different circumstances: (1) when the intermediate is the dominant species (0 M urea) and (2) when the intermediate does not accumulate (2.1 M urea). When Bet v 1 is refolded into buffer in 0 M urea, the two folding scenarios predict that the intermediate outnumbers the denatured state by a factor of 30−50 in the absence of salt prior to folding, whereas in 2.1 M urea, the denatured state outnumbers the intermediate by a factor of 100−200. The data are remarkably unambiguous (Figure 6). In 0 M urea, salt slows refolding 100-fold between 0 and 0.6 M Na_2SO_4 ($m_{Na_2SO_4,0Murea} = -3.25 \pm 0.25$), whereas in 2.1 M urea, refolding is accelerated 400-fold over the same range ($m_{Na_2SO_4,2.1Murea} = 4.41 \pm 0.25$). This trend is exactly as predicted by the off-pathway model, but is very difficult to reconcile with an on-pathway folding model.

The data are analyzed according to both scenarios (eqs 9 and 10) and the results presented in Table 1. Both scenarios lead to the same predictions of stability, urea midpoint denaturation value, and m_{D-N}. The good agreement between kinetic and equilibrium m and ΔG values indicates that no additional folding steps, such as the slow refolding phase, are required to describe our system (cf. Experimental Procedures).

A Single Nondestabilizing Mutation Can Shift Bet v 1 to Two-State Folding Behavior. To improve the fluorescence signal associated with folding, we produced the Y120W mutant. On the basis of the structure of Bet v 1, the Trp was introduced in position 120 (see Figure 1), because the Trp would be expected to be buried in the native state and thus lead to a significant change in fluorescence compared to the denatured state. Equilibrium unfolding of Y120W is accompanied by a 70% reduction in the magnitude of the (large) fluorescence signal, as opposed to the 30% increase in the magnitude of the (small) Tyr fluorescence signal

associated with unfolding of wild-type Bet v 1 (Figure 2). The stability of Y120W is almost identical to that of wild-type Bet v 1 (Table 1). However, there is an important difference in the folding kinetics. Although Y120W also exhibits biphasic refolding kinetics, there is no rollover for the fast phase at low urea concentrations (Figure 5). Instead, we see the V-shaped chevron plot characteristic of proteins folding without an intermediate (*19*). Importantly, the *m* values and calculated stability are in excellent agreement with equilibrium data (Table 2), which are calculated on the basis of a two-state system. This shows that our system is internally consistent. Together, these observations indicate that the Tyr120 → Trp mutation appears to have reduced the extent of accumulation of the compact intermediate to an undetectable level.

The Homologous Protein Mal d 1 Folds as a Two-State System with a Curved Unfolding Limb and a Dramatically Shifted Reaction Coordinate for the Transition State. To extend our analysis of proteins with large cavities, we characterized Mal d 1, which is 57% identical to Bet v 1 in sequence. Like wild-type Bet v 1, it has no Trp residues, and the 10 Tyr residues do not lead to a fluorescence signal which can be used to monitor folding and unfolding (data not shown). Consequently, all characterization is done by CD. The equilibrium denaturation data indicate that the protein is slightly more stable than Bet v 1 (Table 1). This could possibly be attributed to the predicted smaller size of the cavity (D. E. Otzen, unpublished results), although residue substitutions relative to Bet v 1 could just as well explain the difference in stability.

All kinetic data could be fit to a single exponential without drift, and reveal a chevron plot with a completely linear refolding limb and a curved unfolding limb (Figure 5). The curvature in the unfolding limb has been suggested to arise from a shift in the position or compactness of the transition state in response to changes in denaturant concentration for a protein which otherwise folds without folding or unfolding intermediates (*24*). Accordingly, the data were fitted to eq 8. Previous analysis of curved unfolding limbs in other proteins (*43*) suggests that the unfolding rates vary with denaturant concentration in a linear fashion in the *refolding* region of the chevron plot, where the unfolding rates cannot be measured directly by stopped-flow methods. When we extrapolate the unfolding rate from the midpoint of denaturation using the m_u value calculated at the midpoint of denaturation, we obtain kinetic parameters which are in excellent agreement with the equilibrium data. It is noteworthy that the position of the transition state on the reaction coordinate, the β_T parameter, is shifted to 0.49 in Mal d 1 (at the midpoint of denaturation) from 0.86−0.89 in Bet v 1 and Y120W, indicating that the transition state for folding of Mal d 1 is much less compact than that of Bet v 1.

DISCUSSION

Tyr120, an Example of "Negative Gatekeeping". There is good evidence that the intermediate which accumulates during Bet v 1's folding at low denaturant concentrations is off-pathway. While the data in the chevron plot (Figure 5) are compatible with both off- and on-pathway scenarios, the variation of refolding rates as a function of salt concentration (Figure 6) is a strong indication that the intermediate is off the productive folding pathway. The salt-based kinetic data are also compatible with a scenario in which the intermediate is formally on-pathway, but in which the transition state between the intermediate state and the native state is *less* compact than the intermediate (and therefore is less stabilized by salt, so the activation barrier for refolding increases with salt concentration). Even such an esoteric case would indicate that the intermediate must rearrange substantially (i.e., partially unfold) to proceed with folding. In both cases, the intermediate state has characteristics which set it apart from conventional "productive" intermediate states.

The Tyr120 → Trp mutation, originally introduced to provide a better spectroscopic handle on the folding process, turned out to have the added effect of abolishing the accumulation of this intermediate. On the basis of data from protein S6, we recently introduced the term *gatekeeper residues*, that is, residues whose major function is not so much to stabilize the native state as to prevent the formation of misfolded and/or aggregated states of the protein (*7*, *35*). This was based on the observation that various mutations in S6 strengthen the protein's propensity to form a misfolded compact state or its tendency to aggregate during refolding. Thus, the original residues in the wild-type protein reduced this aberrant behavior. In the case of Bet v 1, Tyr120 seems to be a negative gatekeeper, since it actually encourages the transient formation of a misfolded state, which is subsequently suppressed by the Tyr120 → Trp mutation. This is analogous to the suppressor mutations which inhibit inclusion body formation in bacteriophage protein P22, probably by restricting access to misfolded states with marked aggregation propensities (*44*).

It is remarkable that a single conservative mutation can have such dramatic consequences. The strict linearity of Y120W's chevron plot down to 0 M urea implies that the intermediate has been destabilized relative to the denatured state by a factor of at least 100, translating into at least 2.7 kcal/mol. The transition state as such is not stabilized by the mutation relative to the denatured state, since the folding rate for Y120W is actually decreased somewhat compared to the extrapolated value for the wild type (Table 1). Since compact intermediate states generally have fluctuating structure, individual residues should contribute less to stability in an intermediate than in the native state. However, a severe steric clash caused by the extra imidazole ring of Trp, compared to Tyr, has presumably destabilized the intermediate sufficiently to eliminate its accumulation altogether, although there is no comparable effect on the stability of the native state. This is a hallmark of a gatekeeper residue. In contrast, for proteins folding through on-pathway intermediates, it is not unusual for mutations or increases in temperature to have destabilizing effects on both the native state and the intermediate state (*45*, *46*), or on the native state but not the intermediate state (*47*). For example, the destabilizing V26A mutation caused the three-state folder ubiquitin to revert to apparent two-state behavior, although the protein still manifested three-state folding in the presence of the stabilizing Na_2SO_4 salt. We attempted to produce additional mutants of Bet v 1 to ascertain whether introduction of a Trp residue elsewhere in the protein would have a similar effect; unfortunately, these mutants were not expressed in *Escherichia coli* at levels sufficient for biophysical analysis (data not shown).

Partially folded states classified as on-pathway intermediates (according to a D ↔ I ↔ N scheme) can also contain elements of non-native structure, as reported for Im7 (*48*), HPr (*49*), and Rd-apocyt b_{562} (*50*). In these cases, non-native hydrophobic side chain interactions are presumably repaired too rapidly to decrease folding rates; for Im7, they may even be required to maximize stabilization of the folding intermediate (*48*). These effects do not appear to be at play in Bet v 1, where removal of non-native interactions seems to be the key to fast and productive folding. It would seem highly unlikely that the interactions in Bet v 1's C state, which are destabilized by the Y120W mutation, are native-like in origin but lead to a kinetic impasse. If this were the case, the Y120W mutation would also destabilize the native state, and that is clearly not the case. In a biological context, it probably has no serious implications on reducing a folding rate from 125 s^{-1} (Bet v 1 Y120W) to 13 s^{-1} (wild-type Bet v 1). This is presumably why there has been no selection against this residue.

Shift in Transition State Position among Homologous Proteins. We cannot simply attribute the existence of the off-pathway intermediate to the presence of the cavity, since the cavity-containing Mal d 1, which also has a Tyr in the equivalent position, does not accumulate an intermediate. Despite the close sequence homology between Bet v 1 and Mal d 1, the structures of the transition states appear to be very different, with that of Bet v 1 being significantly more compact ($\beta_T = 0.87$) than that of Mal d 1 ($\beta_T = 0.49$) at the midpoint of denaturation. An additional change is the presence of substantial curvature in the unfolding limb of Mal d 1's chevron plot, which implies that the value of β_T ($1 - m_u/m_{D-N}$) changes with denaturant concentration, becoming larger at higher urea concentrations, but still only 0.75 at 5 M urea.

It is a subject of intense debate whether structurally homologous proteins fold by similar pathways (*51*). Irrespective of structural class, the clear trend is for β_T to be preserved within a structural class, although the structure at the residue level may vary significantly due to different local structural propensities for different sequences (*51*, *52*). There is, however, precedence for shifts in β_T at different denaturant concentrations induced by destabilizing mutations in S6 (*24*), the Arc repressor (*53*, *54*), and tendamistat (*55*, *56*). This variation in β_T values can be rationalized as transition state movements over broad and rugged activation barriers, where mutations and changes in denaturant concentration change the position of the highest bump. Alternatively, they could represent a switch from one pathway to another. A third possibility is to regard the phenomenon as a switch between two transition states, separated by a high-energy intermediate which does not accumulate (*55*, *56*). Whatever the reason, it is implicit in these interpretations that the sequence variation between Bet v 1 and Mal d 1 could readily lead to the dramatic difference in folding behavior. At this stage, it is not possible for us to evaluate which of these proposals is the most plausible explanation for the curvature observed in Mal d 1.

Local structural propensity could be an additional reason for the divergence in folding behavior between Bet v 1 and Mal d 1 [cf. Im7 and Im9 (*57*)]. AGADIR (*58*) predicts that Bet v 1's average helical propensity (2.61%) is more than twice as high as that of Mal d 1 (1.11%), and this is even more clearly shown at the level of individual residues constituting the three α-helices. It is tempting to speculate that the reason for Bet v 1's more compact transition state lies in the pronounced helix propensity in the denatured state compared to that of Mal d 1. The following scenario could be proposed. The denatured state in urea at low pH has a very low level of structure (as manifested in the high value of m_{D-N}). However, upon initiation of folding by transfer of Bet v 1 from urea at low pH to a low denaturant concentration at neutral pH, the denatured state rapidly rearranges to a state with significant helical structure, leading in turn to more helical structure in the transition state. Conversely, Mal d 1 has less helical structure in the denatured state at the start, and attains a lower level of structure in the transition state. This implies that the folding pathways of the two proteins are substantially different. Resolution of this issue requires a comprehensive ϕ value analysis based on a large number of mutants (*59*).

Note that local propensity differences cannot explain the formation of an off-pathway intermediate, C. Bet v 1 also exhibits high α-helix propensity around residues 93−102, which constitute part of β-strand 6. However, Mal d 1 has a similar α-helix propensity in the same region (residues 99−108). Furthermore, the Y120W mutation, which abolishes accumulation of C, does not lead to any difference in the helix propensity profile. More complex structural interactions are presumably involved in the formation of C.

ACKNOWLEDGMENT

We thank Kåre Meno for help in preparing Figure 1A.

REFERENCES

1. Levinthal, C. (1968) Are There Pathways for Protein Folding? *J. Chim. Phys. 85*, 44−45.
2. Wetlaufer, D. B. (1973) Nucleation, rapid folding and globular intrachain regions in proteins, *Proc. Natl. Acad. Sci. U.S.A. 70*, 697−701.
3. Ikai, A., and Tanford, C. (1971) Kinetic evidence for incorrectly folded intermediate states in the refolding of denatured proteins, *Nature 230*, 100−102.
4. Tsong, T. Y., Baldwin, R. L., and Elson, E. L. (1971) The sequential unfolding of ribonucleases A: Detection of a fast initial phase in the kinetics of unfolding, *Proc. Natl. Acad. Sci. U.S.A. 68*, 2712−2715.
5. Jackson, S. E., and Fersht, A. R. (1991) Folding of Chymotrypsin Inhibitor 2. 1: Evidence for a two-state transition, *Biochemistry 30*, 10428−10435.
6. Jackson, S. E. (1998) How do small single-domain proteins fold? *Folding Des. 3*, R81−R91.
7. Otzen, D. E., and Oliveberg, M. (1999) Salt-induced detour through compact regions of the protein folding landscape, *Proc. Natl. Acad. Sci. U.S.A. 96*, 11746−11751.
8. Dill, K. A., and Chan, H. S. (1997) From Levinthal to pathways to funnels, *Nat. Struct. Biol. 4*, 10−19.
9. Dill, K. (1999) Polymer principles and protein folding, *Protein Sci. 8*, 1166−1180.
10. Lazaridis, T., and Karplus, M. (1997) "New view" of protein folding reconciled with the old through multiple unfolding simulations, *Science 278*, 1928−1931.
11. Fersht, A. R. (1999) *Structure and mechanism in protein science. A guide to enzyme catalysis and protein folding*, Freeman & Co., New York.
12. Ferguson, N., and Fersht, A. R. (2003) Early events in protein folding, *Curr. Opin. Struct. Biol. 13*, 75−81.
13. Gajhede, M., Osmark, P., Poulsen, F. M., Ipsen, H., Larsen, J. N., van Neerven, R. J. J., Schou, C., Løwenstein, H., and

Spangfort, M. D. (1996) X-ray and NMR structure of Bet v 1, the origin of birch pollen allergy, *Nat. Struct. Biol. 3*, 1040−1045.

14. McCaldon, P., and Argos, P. (1988) Oligopeptide biases in protein sequences and their use in predicting protein coding regions in nucleotide sequences, *Proteins 4*, 99−122.
15. Mogensen, J. E., Wimmer, R., Larsen, J. N., Spangfort, M. D., and Otzen, D. E. (2002) The major birch allergen, Bet v 1, shows affinity for a broad spectrum of physiological ligands, *J. Biol. Chem. 277*, 23684−23692.
16. Markovic-Housley, Z., Degano, M., Lamba, D., von Roepenack-Lahaye, E., Clemens, S., Susani, M., Ferreira, F., Scheiner, O., and Breiteneder, H. (2003) Crystal structure of a hypoallergenic isoform of the major birch pollen allergen Bet v 1 and its likely biological function as a plant steroid carrier, *J. Mol. Biol. 325*, 123−133.
17. Dill, K. A. (1990) Dominant forces in protein folding, *Biochemistry 29*, 7133−7155.
18. Spangfort, M. D., Ipsen, H., Sparholt, S. H., Aasmul-Olsen, S., Larsen, M. R., Mørtz, E., Roepstorff, P., and Larsen, J. N. (1996) Characterization of purified recombinant Bet v 1 with authentic N-terminus, cloned in fusion with maltose-binding protein, *Protein Expression Purif. 8*, 365−373.
19. Tanford, C. (1970) Protein denaturation. Part C. Theoretical models for the mechanism of denaturation, *Adv. Protein Chem. 24*, 1−95.
20. Pace, C. N. (1986) Determination and analysis of urea and guanidine hydrochloride denaturation curves, *Methods Enzymol. 131*, 266−279.
21. Clarke, J., and Fersht, A. R. (1993) Engineered disulfide bonds as probes of the folding pathway of barnase: increasing the stability of proteins against the rate of denaturation, *Biochemistry 32*, 4322−4329.
22. Yadav, S., and Ahmad, F. (2000) A new method for the determination of stability parameters of proteins from their heat-induced denaturation curves, *Anal. Biochem. 283*, 207−213.
23. Otzen, D. E., and Oliveberg, M. (2001) A simple way to measure protein refolding rates in water, *J. Mol. Biol. 313*, 479−483.
24. Otzen, D. E., Kristensen, O., Proctor, M., and Oliveberg, M. (1999) Structural changes in the transition state of protein folding: an alternative interpretation of curved chevron plots, *Biochemistry 38*, 6499−6511.
25. Baldwin, R. (1996) On-pathway versus off-pathway folding intermediates, *Folding Des. 1*, R1−R8.
26. Pace, C. N. (1990) Conformational stability of globular proteins, *Trends Biochem. Sci. 15*, 14−17.
27. Privalov, P. L. (1979) Stability of Proteins, *Adv. Protein Chem. 33*, 167−236.
28. Spolar, R. S., Livingstone, J. R., and Record, M. T. (1992) Use of liquid hydrocarbon and amide transfer data to estimate contributions to thermodynamic functions of protein folding from the removal of nonpolar and polar surface from water, *Biochemistry 31*, 3947−3955.
29. Livingstone, J. R., Spolar, R. S., and Record, M. T. (1991) Contribution to the thermodynamics of protein folding from the reduction in water-accessible nonpolar surface area, *Biochemistry 30*, 4237−4244.
30. Makhatadze, G. I., and Privalov, P. L. (1990) Heat capacity of proteins I. Partial molar heat capacity of individual amino acids in aqueous solution: hydration effect, *J. Mol. Biol. 213*, 375−384.
31. Schindler, T., and Schmid, F. X. (1996) Thermodynamic properties of an extremely rapid protein folding reaction, *Biochemistry 35*, 16833−16842.
32. Myers, J. K., Pace, C. N., and Scholtz, J. M. (1995) Denaturant *m* values and heat capacity changes: Relation to changes in accessible surface areas of protein unfolding, *Protein Sci. 4*, 2138−2148.
33. Privalov, P. L., and Makhatadze, G. I. (1990) Heat Capacity of Proteins II. Partial Molar Heat Capacity of the Unfolded Polypeptide Chains of Proteins: Protein Unfolding Effects, *J. Mol. Biol. 213*, 385−391.
34. Silow, M., and Oliveberg, M. (1997) Transient aggregates in protein folding are easily mistaken for folding intermediates, *Proc. Natl. Acad. Sci. U.S.A. 94*, 6084−6086.
35. Otzen, D. E., Kristensen, P., and Oliveberg, M. (2000) Designed protein tetramer zipped together with an Alzheimer sequence: a structural clue to amyloid assembly, *Proc. Natl. Acad. Sci. U.S.A. 97*, 9907−9912.
36. Kiefhaber, T., Kohler, H.-H., and Schmid, F. X. (1992) Kinetic coupling between protein folding and prolyl isomerisation I. Theoretical Models, *J. Mol. Biol. 224*, 217−229.
37. Kiefhaber, T., Grunert, H. P., Hahn, U., and Schmid, F. X. (1990) Replacement of a cis proline simplifies the mechanism of ribonuclease T1 folding, *Biochemistry 29*, 6475−6480.
38. Jackson, S. E., and Fersht, A. R. (1991) Folding of Chymotrypsin Inhibitor 2. 2: Influence of proline isomerization on the folding kinetics and thermodynamic characterization of the transition state of folding, *Biochemistry 30*, 10436−10443.
39. Nall, B. (1994) in *Mechanisms of protein folding* (Pain, R. H., Ed.) pp 80−103, IRL Press, Oxford, U.K.
40. Pappenberger, G., Aygun, H., Engels, J. W., Reimer, U., Fischer, G., and Kiefhaber, T. (2001) Nonprolyl cis peptide bonds in unfolded proteins cause complex folding kinetics, *Nat. Struct. Biol. 8*, 452−458.
41. Capaldi, A. P., Shastry, M. C., Kleanthous, C., Roder, H., and Radford, S. E. (2001) Ultrarapid mixing experiments reveal that Im7 folds via an on-pathway intermediate, *Nat. Struct. Biol. 8*, 68−72.
42. Timasheff, S. (2002) Protein hydration, thermodynamic binding, and preferential hydration, *Biochemistry 41*, 13473−13482.
43. Otzen, D. E., and Oliveberg, M. (2002) Conformational plasticity in folding of the split β-α-β protein S6: evidence for burst-phase disruption of the native state, *J. Mol. Biol. 317*, 613−627.
44. Mitraki, A., Fane, B., Haase-Pettingell, C., Sturtevant, J., and King, J. (1991) Global suppression of protein folding defects and inclusion body formation, *Science 253*, 54−58.
45. Khorasanizadeh, S., Peters, I. D., and Roder, H. (1996) Evidence for a three-state model of protein folding from kinetic analysis of ubiquitin variants with altered core residues, *Nat. Struct. Biol. 3*, 193−205.
46. Choe, S. E., Matsudaira, P. T., Osterhout, J., Wagner, G., and Shakhnovich, E. I. (1998) Folding kinetics of villin 14T, a protein domain with a central β-sheet and two hydrophobic cores, *Biochemistry 37*, 14508−14518.
47. Sanz, J. M., and Fersht, A. R. (1993) Rationally designing the accumulation of a folding intermediate of barnase by protein engineering, *Biochemistry 32*, 13584−13592.
48. Capaldi, A. P., Kleanthous, C., and Radford, S. E. (2002) Im7 folding mechanism: misfolding on a path to the native state, *Nat. Struct. Biol. 9*, 209−216.
49. Canet, D., Lyon, C. E., Scheek, R. M., Robillard, G. T., Dobson, C. M., Hore, P. J., and van Nuland, N. A. (2003) Rapid formation of non-native contacts during the folding of HPr revealed by real-time photo-CIDNP NMR and stopped-flow fluorescence experiments, *J. Mol. Biol. 330*, 397−407.
50. Feng, H., Takei, J., Litsitz, R., Tjandra, N., and Bai, Y. (2003) Specific non-native hydrophobic interactions in a hidden folding intermediate: Implications for protein folding, *Biochemistry 42*, 12461−12465.
51. Gunasekaran, K., Eyles, S. J., Hagler, A. T., and Gierasch, L. M. (2001) Keeping it in the family: folding studies of related proteins, *Curr. Opin. Struct. Biol. 11*, 83−93.
52. Guerois, R., and Serrano, L. (2000) The SH3-fold family: experimental evidence and prediction of variations in the folding pathways, *J. Mol. Biol. 304*, 967−982.
53. Milla, M. E., Brown, B. M., Waldburger, C. D., and Sauer, R. T. (1995) P22 Arc repressor: transition state properties inferred from mutational effects on the rates of protein unfolding and refolding, *Biochemistry 34*, 13914−13919.
54. Jonsson, T., Waldburger, C. D., and Sauer, T. (1996) Nonlinear free energy relationships in Arc repressor unfolding imply the existence of unstable, native-like folding intermediates, *Biochemistry 35*, 4795−4802.
55. Sánchez, I. E., and Kiefhaber, T. (2003) Evidence for sequential barriers and obligatory intermediates in apparent two-state protein folding, *J. Mol. Biol. 325*, 367−376.
56. Bachmann, A., and Kiefhaber, T. (2001) Apparent two-state tendamistat folding is a sequential process along a defined route, *J. Mol. Biol. 306*, 375−386.
57. Ferguson, N., Capaldi, A. P., James, R., Kleanthous, C., and Radford, S. E. (1999) Rapid folding with and without populated intermediates in the homologous four-helix proteins Im7 and Im9, *J. Mol. Biol. 286*, 1597−1608.

58. Muñoz, V., and Serrano, L. (1997) Development of the multiple sequence approximation within the AGADIR model of α-helix formation: comparison with the Zimm-Bragg and Lifson-Roig formalisms, *Biopolymers 41*, 495−509.
59. Fersht, A. R., Matouschek, A., and Serrano, L. (1992) The Folding of an Enzyme I: Theory of Protein Engineering Analysis of Stability and Pathway of Protein Folding, *J. Mol. Biol. 224*, 771−782.
60. Kleywegt, G. J., and Jonas, T. A. (1994) Detection, delineation, measurement and display of cavities in macromolecular structures, *Acta Crystallogr. D50*, 178−185.
61. DeLano, W. L. (2002) *PyMol*, DeLano Scientific, San Carlos, CA.
62. Corpet, F. (1988) Multiple sequence alignment with hierarchical clustering, *Nucleic Acids Res. 16*, 10881−10890.

BI0358622